

Gender Identification Moderates Stereotype Threat Effects on Women's Math Performance

Toni Schmader

University of Arizona

Received February 22, 2001; revised August 16, 2001; accepted August 17, 2001; published online December 3, 2001

This research applies a social identity perspective to situations of stereotype threat. It was hypothesized that individuals would be more susceptible to the performance-inhibiting effects of stereotype threat to the extent that they are highly identified with the group to which a negative stereotype applies. A quasi-experimental study with male and female college students revealed that individual differences in gender identification (i.e., importance placed on gender identity) moderated the effects of gender identity relevance on women's (but not men's) math performance. When their gender identity was linked to their performance on a math test, women with higher levels of gender identification performed worse than men, but women with lower levels of gender identification performed equally to men. When gender identity was not linked to test performance, women performed equally to men regardless of the importance they placed on gender identity. © 2001 Elsevier Science (USA)

A growing body of evidence suggests that racial and gender differences in testing situations might be created and maintained by stereotypes themselves. Steele and Aronson (1995) demonstrated, for example, that African American students perform more poorly than their White counterparts in testing situations where negative stereotypes about African Americans are relevant. These performance differences do not exist, however, when negative group stereotypes are not relevant to African American students' test performance. Thus, stereotype threat has been characterized as a psychological predicament in which individuals are inhibited from performing to their potential by the recognition that possible failure could confirm a negative stereotype that applies to their in-group and, by extension, to themselves. During the years since the publication of these results, evidence of stereotype threat effects has been found among other stigmatized groups and in other domains. Specifically, negative group stereotypes have been shown to decrease the performance of women on math tests (e.g., Schmader, Johns, Keiffer, Healy, & Fairchild-Ollivier, 2001; Spen-

cer, Steele, & Quinn, 1999), Latinos on intelligence tests (Aronson & Salinas, 1997, as cited in Aronson, Quinn, & Spencer, 1998), white men whose math abilities are compared to those of Asian American men (Aronson et al., 1999), children from low socioeconomic backgrounds in academic testing situations (Croizet & Claire, 1998), the elderly on memory tasks (Levy, 1996), and whites with respect to natural athletic ability (Stone, Lynch, Sjomeling, & Darley, 1999).

Given this diversity of evidence from different research laboratories and with different target groups, there is now a good deal of confidence that stereotype threat has an influence on performance. As with any new area of research, however, many theoretical parameters of this effect have yet to be defined and tested. In particular, we do not have an extensive grasp on whether certain psychological variables make some individuals more or less susceptible to stereotype threat effects. One moderating variable that has been explored in past research is domain identification. Steele (1997; see also Steele & Aronson, 1995) originally theorized, and it was later demonstrated (e.g., Aronson et al., 1999; Stone et al., 1999), that individuals are more susceptible to the debilitating effects of stereotype threat to the extent that they are domain identified. They reasoned that only individuals who care a great deal about performing well in a given domain will feel threatened by the possibility that their poor performance could confirm a broader nega-

I thank Michael Johns, Brian Lickel, and Jeff Stone for their comments on earlier drafts of this article and Steven Spencer for sharing the test items used in his research. I also thank Amy Kiefer, Amy Baesler, and Jennifer White for assisting with data collection.

Address correspondence and reprint requests to Toni Schmader, Department of Psychology, University of Arizona, Tucson, AZ 85721. E-mail: schmader@u.arizona.edu.

tive stereotype that exists about their in-group. Put differently, the domain must be central or important to one's self-definition for the individual to feel threatened by having a negative stereotype applied to him or her in that domain. By the same logic, the centrality or importance of a given social identity might be another variable that would moderate the degree to which one feels threatened in performance situations in which negative stereotypes are relevant to that social identity. In the current research, it is hypothesized that when performance is evaluated in light of a social identity that is negatively stereotyped, one's ability to perform well will be inhibited to the extent that one is highly identified with that social group.

Social Identity Theory as Applied to Stereotype Threat

Social identity theory (Tajfel, 1981; Tajfel & Turner, 1986) and self-categorization theory (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987) maintain that individuals possess two sources of identity: their personal identities as unique individuals and various social identities as members of social groups. A basic tenet of social identity theory is that people are motivated to maintain positive social identities in the same way as they are motivated to maintain positive personal identities. Furthermore, just as one's personal identity is threatened when one compares unfavorably to other individuals, a given social identity is threatened when one's in-group compares unfavorably to out-groups. From this perspective, negative group stereotypes are implicit threats to one's social identity simply because such stereotypes often define one group as inferior to another in a given domain.

While social identity theory has not traditionally been applied to performance effects, several stereotype threat studies can be construed in terms of social identity theory. In some research, for example, stereotype threat has been induced in individuals by priming a negative social identity that is relevant to the task. Shih, Pittinsky, and Ambady (1999) primed different social identities in an initial and ostensibly unrelated questionnaire to test the effects of automatic identity activation on subsequent math performance. Among a sample of Asian American women, test performance was higher after priming their Asian identity (i.e., a social identity that is positively stereotyped in math) and lower after priming their gender identity (i.e., a social identity that is negatively stereotyped in math). Similarly, Steele and Aronson (1995, Study 4) found that African American students performed more poorly than Whites when they had been asked to indicate their race on a demographic sheet prior to taking the test but performed equally to Whites when they had not been asked to identify their race. Thus, merely priming one's social identity (and presumably the associated stereotype) before taking a stereotype relevant test leads one to perform in a stereotype-consistent manner.

In addition to merely priming a given social identity, other stereotype threat research has used more explicit manipulations of social identity threat in which it is clearly stated that one's performance will be evaluated with respect to a given social identity one possesses. For example, Aronson et al. (1999) found that White males performed more poorly on a math test if they were told that the researcher would be comparing Asian American and White students' math performance. More than just making the stereotype salient, this manipulation includes an explicit suggestion that one's social identity is "on the line" in the performance situation. Similarly, Schmader et al. (2001) found that women performed significantly worse on a math test when they were told that their performance would be used as an indicator of their personal math ability as well as an indicator of women's math ability in general, as compared to a condition in which their performance was linked only to their personal ability and no mention of gender was made. In other words, women's performance was lower when both their personal and social identities would be implicated by their performance on a math test than when only their personal identity was on the line.

Group Identification as a Moderator of Stereotype Threat Effects

The research reviewed above provides evidence that stereotype threat can occur as a function of either making a negative social identity salient or linking that identity explicitly to one's performance on a task. This research takes a categorical approach to social identity in that it compares the performance of individuals who possess a negative social identity (e.g., women, African Americans) to those who do not (e.g., men, Whites). Although self-categorization theory (Turner et al., 1987) construes social identity in categorical terms, other research maintains that any given social identity might be a more important source of identity for some members of a group than for others (e.g., Brewer & Silver, 2000; Ellemers, Spears, & Doosje, 1997). For example, Luhtanen and Crocker (1992) included group importance as part of their measure of collective self-esteem. Similarly, Smith and his colleagues discussed individual differences in group attachment (Smith, Murphy, & Coats, 1999) and the extent to which one's in-group is included in one's self-concept (Smith & Henry, 1996). Furthermore, research focused specifically on stigmatized groups has examined individual differences in the importance placed on ethnic group membership (e.g., Phinney & Devich-Navarro, 1997; Sellers, Rowley, Chavous, Shelton, & Smith, 1997) and on gender identity (e.g., Burn, Aboud, & Moyles, 2000). Thus, while all women might recognize their membership in the social category "women," there is likely to be variation in the extent to which they consider this category membership to be a central or important part of their self-identity.

One implication of individual differences in group identification is that individuals who consider a given group membership to be an important source of identity should have a stronger motivation to maintain a positive image of that identity and thus should experience greater threat at the suggestion that their in-group is somehow inferior to other groups. Indeed, the importance of a given social identity has been shown to influence how people respond to negative group outcomes in terms of whether or not they attribute those outcomes to discrimination (Major, Quinton, & Schmader, 2001), what kind of action they take (Ellemers et al., 1997), how they evaluate themselves (McFarland & Buehler, 1995), and the extent to which they self-stereotype (Spears, Doosje, & Ellemers, 1997; Verkuyten & Nekuee, 1999). When faced with threats to their social identity, individuals who are highly identified with their group are more likely than those who are less identified to engage in behavioral and psychological strategies designed to protect and maintain that social identity.

Similarly, to the extent that negative stereotypes can inhibit performance because they present a threat to social identity (e.g., Schmader et al., 2001), the impact of these stereotypes on performance should be stronger to the extent that one is highly identified with that group. The primary purpose of the current research was to test group identification as a moderator of stereotype threat effects when that social identity is implicated by one's performance at a stereotype relevant task. This question was examined among a sample of male and female college students taking a difficult math exam. I hypothesized that when their gender identity was linked to their performance on a math test, women would perform more poorly to the extent that they perceive gender to be important to their self-definition. For men, associating their gender identity to their test performance, regardless of the perceived importance of that identity, was not expected to affect their math performance.

METHOD

Participants and Design

Participants were 33 Caucasian male and 32 Caucasian female undergraduates who participated in partial fulfillment of a course requirement.¹ All participants had quantitative scores on the Scholastic Achievement Test (SAT) that fell between 500 and 700 (as reported during an earlier mass testing session). This selection criterion was based on selection standards used in previous work on stereotype threat among women taking math tests (Spencer et al., 1999). Participants were randomly assigned to one of two condi-

tions in which their gender identity was either linked to their performance on a math test or not. Individual differences in the importance of gender identity were examined as a continuous variable in the design. Thus, the complete design of the study included gender of participant, a gender identity relevance manipulation, and the importance of gender identity as predictors of math performance.

Individual Differences in Gender Identification

During a mass testing session at the beginning of the semester, participants completed the importance subscale of the Collective Self-Esteem Scale (Luhtanen & Crocker, 1992) modified to assess the perceived importance of gender identity to self-definition. They rated the following four items (worded specific to their own gender) on a scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*): "Being a woman/man is an important part of my self-image," "Being a woman/man is unimportant to my sense of what kind of person I am" (reverse scored), "Being a woman/man is an important reflection of who I am," and "Being a woman/man has very little to do with how I feel about myself" (reverse scored). Participants' responses to the four items were averaged to form a reliable index of gender identification ($\alpha = .70$). Within the experimental sample, there were no gender differences in gender identification (women: $M = 3.82$, range = 2.00–5.00; men: $M = 3.64$, range = 2.25–5.00), $t(63) = -.94$, $p > .05$. Thus, both men and women tended to view their gender identity as somewhat important to them, although scores did range from very high levels of gender identification to much lower levels of gender identification.

Procedures

Participants arrived at the laboratory in groups of 1 to 4 people and were greeted by a female research assistant, who led them to separate cubicles. Additional instructions, including the experimental manipulation of gender identity relevance, were presented to participants in a tape-recorded message played over an intercom system. Specifically, participants heard the voice of a man who introduced himself as a researcher who was developing a standardized math exam. In both identity relevance conditions, he explained that he was interested in each individual's performance on the test and that he would be comparing participants' individual scores to those of other students. He said that he would be using their scores on the test as an indicator of their personal math ability. Consistent with these instructions, participants were asked to write their first initials and last names on the cover sheets of their tests. In the Gender Identity Not Relevant condition, no further information about the purpose of the research was given, and gender was not mentioned.

In the Gender Identity Relevant condition, the male re-

¹ Only Caucasian participants were recruited based on evidence that the gender gap in math performance exists only in Caucasian samples and not in Hispanic, African American, or Asian American samples (Hyde, Fennema, & Lamon, 1990).

TABLE 1

Hierarchical Regression Analyses Predicting Math Performance from Gender, Gender Identification, and Gender Identity Relevance

	Number correct		Number attempted		Accuracy	
	β	ΔR^2	β	ΔR^2	β	ΔR^2
Step 1		.082*		.113**		.264***
SAT-Q	.29*		-.34**		.53***	
Step 2		.220*		.137		.100
SAT-Q	.23*		-.37**		.51***	
Gender ^a	-.26		.11		-.30*	
Identity Relevance ^b	-.20		.03		-.15	
Group Importance	.07		-.12		-.11	
Gender \times Identity	-.05		-.32		.10	
Gender \times Importance	.13		.12		-.04	
Identity \times Importance	.19		.23		-.11	
Threeway Interaction	-.42*		-.40 ⁺		-.08	

^a Coded as 0 = men, 1 = women.^b Coded as 0 = Gender Identity Not Relevant, 1 = Gender Identity Relevant.⁺ $p < .10$.* $p < .05$.** $p < .01$.*** $p < .001$.

searcher on the tape went on to explain that he was also interested in how women score on the test relative to men. He stated that because he would be comparing women's scores to men's scores, he would be using each individual's score as an indicator of women's or men's math ability in general. In this condition, participants identified their gender on the cover sheets of their tests in addition to providing their last names. Thus, in both gender identity relevance conditions, individuals were made to feel personally invested in their performance. The difference between conditions was whether their gender identity would also be implicated by their performance on the test.

Before beginning the test, participants completed a brief measure of anxiety adapted from Spielberger's State-Trait Anxiety Inventory (Spielberger, Gorsuch, & Lushene, 1970). Participants rated, on a 4-point scale (ranging from *not at all* to *very much*), the extent to which they were feeling worried, nervous, jittery, at ease, relaxed, and comfortable. Responses to positive emotions were reverse scored and averaged with responses to the negative items to create a reliable index of pretest anxiety ($\alpha = .81$). Standard instructions for completing the test were then provided for participants over the intercom. The test consisted of 20 multiple-choice questions drawn from practice tests for the quantitative section of the Graduate Record Examination. All of the questions were word problems requiring advanced algebraic calculations to correctly determine which of the five possible answers was correct. This type of exam is particularly suitable for this research given the results of a previous meta-analysis indicating that gender differences in math performance exist only on tests requiring complex problem solving (Hyde, Fennema, & Lamon, 1990). Participants

were given 20 min to work on the test. After completing the test, participants were thoroughly debriefed about the true purpose of the study and were thanked for their participation.

RESULTS

As expected with random assignment, there was no difference in gender identification between experimental conditions, all F 's < 1.0 , and gender identification was not correlated with participants' quantitative SAT scores, $r = .14$, $p > .10$. Likewise, there were no differences between experimental conditions or between genders in participants' SAT scores, all F 's < 1.0 . Performance scores were analyzed using hierarchical multiple regression. Participants' quantitative SAT scores were entered as a covariate on the first step of the analysis. The second step included three main effects (dummy-coded variables for gender and gender identity relevance and the continuous measure of gender identification) and all possible interactions between those variables. Results of these analyses are summarized in Table 1.

Number Correct

As predicted, analysis of the number of items answered correctly during the 20-min test yielded a significant three-way interaction among gender, gender identity relevance, and gender identification, $\beta = -.42$, $p < .05$ (see Fig. 1). Quantitative SAT scores were a significant predictor of performance on the test, $\beta = .29$, $p < .05$. No other main effects or interactions were significant, all p 's $> .10$. Using procedures outlined by Aiken and West (1991), the inter-

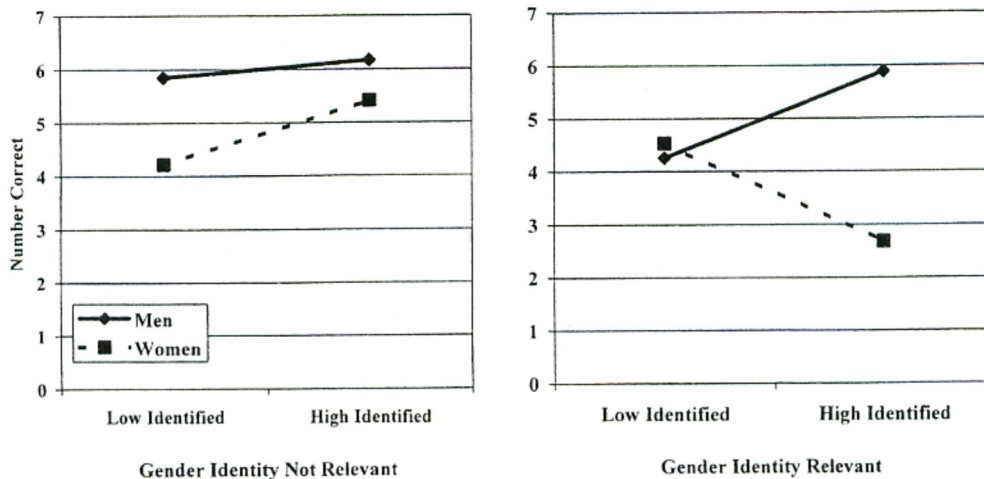


FIG. 1. Gender identification as a moderator of gender identity relevance effects on women's and men's math performance.

action was examined using simple slopes analyses. The first set of analyses tested the simple slopes representing gender differences in performance within each identity relevance condition and evaluated at 1 standard deviation above and below the mean of gender identification ($M = 3.73$, $SD = 0.75$). These analyses revealed that only when gender identity was linked to test performance and gender identification was high did women score significantly lower than men on the test, $\beta = -.69$, $p < .01$. At lower levels of gender identification, women performed comparably to men, even though gender was made explicitly relevant in the testing situation, $\beta = .06$, $p > .10$. Similarly, when gender was not relevant to test performance, women's and men's math performances were equivalent regardless of whether they were high, $\beta = -.16$, $p > .10$, or low, $\beta = -.35$, $p > .10$, in gender identification.

A second set of simple slopes analyses tested the effect of the experimental manipulation separately among men and women and evaluated at high and low levels of gender identification. These analyses revealed that the identity relevance manipulation affected only women who tended to be highly identified with their gender, $\beta = -.59$, $p < .01$. As predicted, highly gender-identified women answered fewer problems correctly when their gender was relevant to the test than when it was not relevant to the test. This same manipulation had no significant effect on women with lower levels of gender identification, $\beta = .07$, $p > .10$, or on men with either high, $\beta = -.06$, $p > .10$, or low levels of gender identification, $\beta = -.34$, $p > .10$.

A final set of analyses tested the simple slopes of gender identification predicting performance within each gender and identity relevance condition. Although none of these simple slopes reached conventional levels of statistical significance, there was a nonsignificant trend for higher levels of gender identification to predict poorer performance among women when gender was relevant to the test, $\beta = -.39$, $p < .12$, but to predict better performance among

men when gender was relevant to the test, $\beta = .35$, $p < .15$. When gender was not relevant to the test, gender identification was unrelated to performance among both men, $\beta = .07$, $p > .25$, and women, $\beta = .25$, $p > .25$.

Number of Items Attempted

Only 1 participant had a chance to attempt all 20 items on the test during the 20-min time period. Although differences in the number of items attempted were not predicted, analysis of this measure revealed a marginally significant three-way interaction, $\beta = -.40$, $p < .10$. Quantitative SAT scores were a significant predictor of the number of items attempted, $\beta = -.34$, $p < .01$, but no other main effects or interactions were significant, all p 's $> .10$. Simple slopes analyses revealed that the marginal three-way interaction was driven by the relatively low number of items attempted by more highly gender-identified women whose gender identity was linked to their test performance. Mirroring the pattern observed for number of items correct, women attempted significantly fewer items than did men when they had higher levels of gender identification and were in the Gender Identity Relevant condition, $\beta = -.61$, $p < .01$; all other simple slopes representing the gender effect were nonsignificant, all β 's $< .20$, all p 's $> .10$. Similarly, making gender identity relevant (rather than not relevant) led women to attempt fewer problems only if they had higher levels of gender identification, $\beta = -.60$, $p < .01$. The identity relevance manipulation had no effect on the number of attempts made by less identified women or by men who were either low or high in gender identification, all β 's $< .20$, all p 's $> .10$.

Accuracy: Percentage Correct Out of Attempts

Because there seemed to be systematic variation in the number of items attempted, the accuracy of participants'

performance was also analyzed. Accuracy was defined as the percentage of items participants answered correctly out of the total number of items they attempted (Steele & Aronson, 1995). Analysis of participants' accuracy in their responses revealed a main effect only of gender, $\beta = -.31$, $p = .05$. Men were more accurate in their answers than were women. Quantitative SAT scores were a significant covariate in the analysis, $\beta = .51$, $p < .001$, but no other main effects or interactions were significant, all p 's $> .10$. Thus, gender identified women appeared to be answering fewer items correctly on the test when their gender identity was linked to their performance because they tended to attempt fewer items overall and not because they were significantly less accurate in the questions they did attempt to answer.

Anxiety

Previous research has found evidence that anxiety mediates stereotype threat effects experienced by women while taking a math test (Spencer et al., 1999). In the current research, however, regression analyses revealed no significant main effects or interactions on participants' self-reported anxiety just prior to taking the test, all p 's $> .10$. Overall, participants' anxiety scores were neither extremely low nor extremely high (grand mean = 2.03 on a 4-point scale). Thus, although women's math performance was affected by having their gender identity linked to the test (assuming that identity was an important aspect of their self-definition), this performance impairment was not paralleled by self-reported feelings of anxiety.

DISCUSSION

Social identity theory posits that individuals are motivated to maintain a positive sense of social identity. For any given social identity, however, this motivation is stronger among those who feel that that identity is an important aspect of their self-definition. When that social identity is subject to scrutiny through the lens of a negative stereotype, those who are highly identified with their social group should experience the greatest degree of stereotype threat and resulting impairments to their performance. Results of the current research are consistent with this hypothesis. Women showed poorer performance compared to men on a stereotype relevant task when their social identity was linked to their test performance, but only if they considered gender to be an important part of their self-definition. Women who did not feel that gender was central to their self-concept performed equally to men regardless of a manipulation that might be expected to produce stereotype threat in any women identified with their math performance.

It is also important to note that performance is not inhibited by having any important social identity linked to per-

formance in a testing situation. If this were the case, then men might have also shown a decrement in performance to the extent that their gender identity was important to them and was linked to their test performance. If anything, when gender was relevant to the test, men showed a slight tendency to perform better to the extent that gender was an important source of their identity. Contrast this pattern to women who performed slightly worse to the extent that their gender was an important identity that was on the line in the testing situation.

These results add to a growing body of work highlighting the role of identity structure in moderating the performance effects of negative group stereotypes. For example, while other research suggests that the importance placed on the *domain* moderates stereotype threat effects (Aronson et al., 1999), the results presented here suggest that the importance of group membership can also moderate these effects to the extent that individuals feel that their performance will be taken as representative of group ability. Prior research using a stereotype threat framework to examine gender differences in math performance has not taken into account individual differences in gender identification. In general, this past research finds that, controlling for individual variation in ability, women do more poorly than men on challenging tests that are described as diagnostic of math ability (e.g., Quinn & Spencer, 2001; Spencer et al., 1999). Although there was a hint of such gender differences in performance in the current data, the significant three-way interaction suggests that additional variance in performance can be explained by examining the moderating effects of situational cues to identity relevance and individual differences in gender identification. Women who did not see gender as an important social identity did not show lower performance under conditions of stereotype threat. Thus, merely possessing a negative social identity is not sufficient to make one susceptible to situations that produce stereotype threat. By the same token, high levels of gender identification do not predict lower math performance among women regardless of situational context. Rather, it is the interaction of situational factors with features of self-definition that best predicts lowered performance.

Underlying Mechanisms and Future Directions

While the results of this study suggest that group identification might be an important moderator of stereotype threat effects, the process by which these effects occur remains less clear. On the one hand, these results are consistent with a social identity perspective arguing that gender identified women experience greater threat in situations where their gender identity is linked to their math performance. However, we did not see differences in pretest anxiety that correspond to the performance differences. Without this direct evidence of threat, alternative explanations for the process underlying these effects should be

considered. One intriguing possibility is that the performance differences observed might be due to a more general automatic behavior effect of primed stereotypes rather than to a fear among women that they might confirm a stereotype about their gender (for a review, see Wheeler & Petty, in press). Research on automatic behavior effects has found that activating a stereotype can lead individuals to behave in a way that is stereotype consistent, even though they themselves are not members of the group to which the stereotype applies (Dijksterhuis, Aarts, Bargh, & van Knippenberg, 2000; Dijksterhuis & van Knippenberg, 1998). Furthermore, these studies show that the strength of the behavior is linearly related to the strength of the association between the group and the stereotype. In the current study, when gender was explicitly linked to the test, both women's and men's performance tended to be consistent with the stereotype of their gender (i.e., women did slightly worse and men did slightly better) to the extent that gender was an important part of their identity. Thus, gender identification might have increased participants' sensitivity to an automatic stereotype-behavior link, perhaps because those who are gender identified have a stronger association between gender and math stereotypes.

Future research will need to examine the distinction between stereotype threat effects and automatic behavior effects. Building off of the current research, group identification should be examined as a moderator in automatic behavior paradigms where stereotypes are implicitly primed for individuals who are or are not targeted by those stereotypes. In one recent study (Wheeler, Jarvis, & Petty, 2001), for example, Whites performed more poorly on a test if they were first asked to write about a day in the life of "Tyrone Walker," a manipulation that was meant to prime African American stereotypes. Interestingly, however, this effect was stronger among individuals who wrote their story from a first-person viewpoint. This finding suggests that automatic behavior effects can be moderated by the extent to which the stereotype or group prime is referenced to oneself (Wheeler & Petty, in press). Thus, identification might not need to be constrained by group membership but could represent the degree to which any primed stereotype or identity becomes associated to one's concept of self. Men who self-identify with a feminine identity might perform more poorly on a math test that implicates that identity, whereas women who self-identify with a masculine identity might perform better. Identification processes, therefore, might be an important moderator of effects that occur through either a threat-based or an automatic behavior mechanism.

Regardless of the underlying mechanism, the results of this research suggest that individuals who derive a strong sense of identity from a negatively stereotyped group might be more at risk for lowered performance when negative stereotypes about that group are relevant in testing situa-

tions. It is interesting to consider these findings along with recent evidence that one's level of identification with a negatively stereotyped group has benefits for self-esteem (Branscombe, Schmitt, & Harvey, 1999). These self-esteem benefits of in-group identification might in some situations come at a cost of lower performance. Likewise, individuals who are highly motivated to achieve in a domain in which their group is negatively stereotyped might actively reduce the importance of that social identity as a way to alleviate future experiences of stereotype threat. Evidence that individuals distance themselves from stereotypical preferences after experiencing stereotype threat might be suggestive of this sort of identity negotiation (Steele & Aronson, 1995), although this hypothesis merits additional investigation.

Finally, in the current study, we found performance effects on the absolute number of items answered correctly and, to a lesser degree, on the number of items attempted. Although effects on the number of items attempted were only marginally significant in this study, this pattern has also been found in other stereotype threat studies (e.g., Steele & Aronson, 1995). These results might suggest that women who were gender identified worked more slowly on the test when they knew they would be evaluated as women. However, an effort to be careful or thorough on any given problem translated into fewer problems answered correctly. This interpretation assumes that women who feel they must act as representatives of their gender are motivated to perform better on that task than women who do not feel as closely linked to their gender identity. Other possibilities are that the greater amount of time spent on each problem is a result of greater mental distraction due to the relevance of a negative stereotype about a group that is highly important to one's self-definition or reduced motivation perhaps stemming from lowered expectancies of good performance.

Additional research on the mediating mechanisms of stereotype threat phenomena is still needed, and in particular, future work should attempt to isolate the threat-based versus automatic association explanation for performance effects found in stereotype threat studies. Understanding the underlying process at work will be necessary to create solutions for the group performance differences observed because of the existence of negative social stereotypes.

REFERENCES

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999). When White men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology*, 35, 29-46.
- Aronson, J., Quinn, D. M., & Spencer, S. J. (1998). Stereotype threat and the academic underperformance of minorities and women. In J. K. Swim & C. Stangor (Eds.), *Prejudice: The target's perspective* (pp. 83-103). San Diego: Academic Press.

- Branscombe, N. R., Schmitt, M. T., & Harvey, R. D. (1999). Perceiving pervasive discrimination among African Americans: Implications for group identification and well-being. *Journal of Personality and Social Psychology*, *77*, 135-149.
- Brewer, M. B., & Silver, M. D. (2000). Group distinctiveness, social identification, and collective mobilization. In S. Stryker & T. J. Owens (Eds.), *Self, identity, and social movements* (pp. 153-171). Minneapolis: University of Minnesota Press.
- Burn, S. M., Aboud, R., & Moyles, C. (2000). The relationship between gender social identity and support for feminism. *Sex Roles*, *42*, 11-12.
- Croizet, J. C., & Claire, T. (1998). Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin*, *24*, 588-594.
- Dijksterhuis, A., Aarts, H., Bargh, J. A., & van Knippenberg, A. (2000). On the relation between associative strength and automatic behavior. *Journal of Experimental Social Psychology*, *36*, 531-544.
- Dijksterhuis, A., & van Knippenberg, A. (1998). The relation between perception and behavior, or how to win a game of Trivial Pursuit. *Journal of Personality and Social Psychology*, *74*, 865-877.
- Ellemers, N., Spears, R., & Doosje, B. (1997). Sticking together or falling apart: Ingroup identification as a psychological determinant of group commitment versus individual mobility. *Journal of Personality and Social Psychology*, *72*, 617-626.
- Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, *107*, 139-155.
- Levy, B. (1996). Improving memory in old age through implicit self-stereotyping. *Journal of Personality and Social Psychology*, *71*, 1092-1107.
- Luhtanen, R., & Crocker, J. (1992). A collective self-esteem scale: Self-evaluation of one's social identity. *Personality and Social Psychology Bulletin*, *18*, 302-318.
- Major, B., Quinton, W., & Schmader, T. (2001). *Reconsidering the antecedents and consequences of attributions to discrimination*. Unpublished manuscript, University of California, Santa Barbara.
- McFarland, C., & Buehler, R. (1995). Collective self-esteem as a moderator of the frog-pond effect in reactions to performance feedback. *Journal of Personality and Social Psychology*, *68*, 1055-1070.
- Phinney, J. S., & Devich-Navarro, M. (1997). Variations in bicultural identification among African American and Mexican American adolescents. *Journal of Research on Adolescence*, *7*, 3-32.
- Quinn, D. M., & Spencer, S. J. (2001). The interference of stereotype threat with women's generation of mathematical problem-solving strategies. *Journal of Social Issues*, *57*, 55-71.
- Schmader, T., Johns, M., Keiffer, V. A., Healy, E. A., & Fairchild-Ollivierre, S. (2001). *Stereotype threat as "double devaluation"*. Unpublished manuscript. University of Arizona.
- Sellers, R. M., Rowley, S. A. J., Chavous, T. M., Shelton, J. N., & Smith, M. A. (1997). Multidimensional inventory of Black identity: A preliminary investigation of reliability and construct validity. *Journal of Personality and Social Psychology*, *73*, 805-815.
- Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science*, *10*, 80-83.
- Smith, E. R., & Henry, S. (1996). An in-group becomes part of the self: Response time evidence. *Personality and Social Psychology Bulletin*, *22*, 635-642.
- Smith, E. R., Murphy, J., & Coats, S. (1999). Attachment to groups: Theory and measurement. *Journal of Personality and Social Psychology*, *77*, 94-110.
- Spears, R., Doosje, B., & Ellemers, N. (1997). Self-stereotyping in the face of threats to group status and distinctiveness: The role of group identification. *Journal of Personality and Social Psychology*, *23*, 538-553.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*, 4-28.
- Spielberger, C. D., Gorsuch, R. R., & Lushene, R. (1970). *The State-Trait Anxiety Inventory (STAI) test manual*. Palo Alto, CA: Consulting Psychologists Press.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape the intellectual identities and performance. *American Psychologist*, *52*, 613-629.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797-811.
- Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on Black and White athletic performance. *Journal of Personality and Social Psychology*, *77*, 1213-1227.
- Tajfel, H. (1981). *Human groups and social categories: Studies in social psychology*. Cambridge, UK: Cambridge Univ. Press.
- Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *The psychology of intergroup relations* (pp. 7-24). Chicago: Nelson-Hall.
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization*. Oxford, UK: Blackwell.
- Verkuyten, M., & Nekuee, S. (1999). Ingroup bias: The effect of self-stereotyping, identification, and group threat. *European Journal of Social Psychology*, *29*, 411-418.
- Wheeler, S. C., Jarvis, W. B. G., & Petty, R. E. (2001). Think unto others: The self-destructive impact of negative racial stereotypes. *Journal of Experimental Social Psychology*, *37*, 173-180.
- Wheeler, S. C., & Petty, R. E. (in press). The effects of stereotype activation on behavior: A review of possible mechanisms. *Psychological Bulletin*.